Naiomi T. Cameron
cameron@math.hmc.edu

STATEMENT OF RESEARCH

The sequence of numbers known as the Catalan numbers $1, 1, 2, 5, 14, 42, \ldots, \frac{1}{n+1}\binom{2n}{n}, \ldots$ are without question one of the most important sequences in all of enumerative combinatorics. Richard Stanley cites at least 66 combinatorial settings where the sequence appears [7]. Among the numerous interpretations of the Catalan numbers, we have the number of paths restricted to the first quadrant of the x, y-plane starting at $(0, 0)$ and ending at $(2n, 0)$ using "up" steps $(1, 1)$ or "down" steps $(1, -1)$ (known as Dyck paths). This research will focus on a generalization of the Catalan numbers which can be interpreted as taking Dyck paths and perturbing the length of the down step. One particular example of this generalization is given by the sequence of numbers known as the ternary numbers,

$$1, 1, 3, 12, 55, 273, 1428, 7752, 43263, \ldots, \frac{1}{2n+1}\binom{3n}{n}, \ldots.$$

Although this sequence does not have a tridiagonal Stieltjes matrix, it arises in many natural contexts and extensions of known results. Hence, we use this sequence as an important example of something which lies on the boundary of what is known and what is new. Much of this research is an attempt to extend many of the known results for the Catalan numbers to ternary (m-ary) numbers. There are many combinatorial objects counted by the ternary numbers. Some of these objects are:

- paths in the first quadrant starting at $(0, 0)$ and ending at $(3n, 0)$ using steps $\{(1, 1), (1, -2)\}$,
- paths in the first quadrant starting at $(0, 0)$ and ending at $(2n, 0)$ using steps

  $\{(1, 1), (1, -1), (1, -3), (1, -5), (1, -7), \ldots, \}$,
- rooted planar trees with $2n$ edges where every node (including root) has even outdegree,
- rooted planar trees with $3n$ edges where every node (including root) has outdegree zero or three,
- sequences $i_1 i_2 i_3 \cdots i_{2n}$ such that (i) $i_k + 1$ is even for $k = 1, 2, \ldots, 2n$, (ii) $i_1 + i_2 + \ldots + i_k \geq 0$ for $k = 1, \ldots, 2n$, and (iii) $i_1 + i_2 + \ldots + i_{2n} = 0$,
- diagonally convex directed polyminoes with n diagonals.

A list of bijections between these objects is included in chapter 6 of [7].

The Riordan group is used extensively throughout this research as a combinatorial tool for solving enumeration problems involving random walks and trees. As the recent work of Sprugnoli et al. demonstrates, there are also many interesting questions to be answered about structure in the Riordan group itself, a few which are explored here. Another aspect of this research is to develop new tools in the study of random walks; in particular, to extend the use of Riordan group techniques, and to study the Riordan group (and other structures which generalize to the Riordan group) in search of combinatorial interpretations of sub-structures, such as (pseudo) involutions.

Definition 0.1.1. *An infinite lower triangular matrix,* $L = (l_{n,k})_{n,k \geq 0}$ *is a* **Riordan matrix** *if there* *exists generating functions* $g(z) = \sum g_n z^n$, $f(z) = \sum f_n z^n$, $f_0 = 0$, $f_1 \neq 0$ *such that* $l_{n,0} = g_n$ *and* $\sum_{n\geq k} l_{n,k} z^n = g(z)\left(f(z)\right)^k$.

From the definition it is clear that a Riordan matrix L is completely defined by the functions $g(z)$ and $f(z)$, hence if L is called Riordan and we write $L = (g(z), f(z))$. We can define multiplication of Riordan matrices in the following way. Let $L_1 = (g_1(z), f_1(z))$ and $L_2 = (g_2(z), f_2(z))$ then

$$L_1 L_2 = (g_1(z) g_2(f_1(z)), f_2(f_1(z)))$$

The set of Riordan matrices along with this multiplication gives us a group which we refer to as the Riordan group.

§1 INTERPRETATIONS OF THE TERNARY NUMBERS

1.1 GENERALIZED DYCK PATHS

Let $\mathcal{T}$ denote the set of all lattice paths on $Z \times Z$ which start at the origin $(0,0)$, stay in the first quadrant, and use either a $(1,1)$ "up" step or $(1,-2)$ "down" step. Let $\mathcal{T}_{n,k}$ denote the paths in $\mathcal{T}$ ending at $(3n, k)$, so that $\mathcal{T} = \bigcup_{n,k\geq 0} \mathcal{T}_{n,k}$. We will refer to paths in $\mathcal{T}_{n,0}$ as generalized Dyck paths. Let $t_{n,k} = |\mathcal{T}_{n,k}|$ and $t_{n,0} = t_n$.

Theorem 1.1.1. *Let* $T(z) = \sum_{n=0}^{\infty} t_n z^n$. *Then:*

(i.) $T(z) = 1 + zT^3(z)$,

(ii.) $t_{n,k} = [z^n] z^k T^k(z)$.

Corollary 1.1.2.

(i.) $t_n = \frac{1}{2n+1}\binom{3n}{n}$

(ii.) $t_{n,k} = \frac{k}{3n+k}\binom{3n+k}{n}$.

1.2 STATISTICAL RESULTS

We would like to make some statistical inferences about the paths in $\mathcal{T}$. Let us strict $\mathcal{T}$ to those paths which have terminal height 0 and subcount by number of returns to the x-axis. Denote by $r_{n,j}$ the number of paths of length n with j returns to the x-axis.

$$R = (r_{n,j})_{n,j\geq 0} = \begin{bmatrix} 1 & & & & & & \\ 0 & 1 & & & & & \\ 0 & 2 & 1 & & & & \\ 0 & 7 & 4 & 1 & & & \\ 0 & 30 & 18 & 6 & 1 & & \\ 0 & 143 & 88 & 33 & 8 & 1 & \\ 0 & 728 & 455 & 182 & 52 & 10 & 1 \end{bmatrix} = \left(1, T^2(z)\right)$$

One might ask, on average, how many returns to the x-axis can we expect to see? To answer this, we calculate the associated probabilities.

Proposition 1.2.1. *Let $p_m(n)$ denote the probability that a randomly chosen path of length n has m returns. Then*

$$p_m(n) = \frac{2m(2n+1)(n-m+1)(n-m+2)\cdots(n-1)}{3(3n-m)(3n-m+1)(3n-m+2)\cdots(3n-1)}$$

It follows immediately that $p_m(n) \to \frac{4m}{3^{m+1}}$ as $n \to \infty$.

Proposition 1.2.2. *Let $Y(n)$ denote the random variable for the number of returns to the x-axis. Define $P_{Y(n)}(z) := \sum_{m=0}^{\infty} p_m(n) z^n$. Then*

$$P_{Y(n)}(z) = \frac{2(2n+1)}{3(3n-1)} zF(2, 1-n; 2-3n; z)$$

where $F(a_1, a_2, \ldots, a_n; b_1, b_2, \ldots, b_m; z)$ is a hypergeometric function with upper parameters $a_1, \ldots, a_n$ and lower parameters $b_1, \ldots, b_m$ of the form

$$F(a_1, a_2, \ldots, a_n; b_1, b_2, \ldots, b_m; z) = \sum_{j=0}^{\infty} \frac{a_1^{\bar{j}} a_2^{\bar{j}} \cdots a_n^{\bar{j}}}{b_1^{\bar{j}} b_2^{\bar{j}} \cdots b_m^{\bar{j}}} \cdot \frac{z^j}{j!}$$

where $x^{\bar{j}} = x(x+1)(x+2)\cdots(x+j-1)$. It follows from the definition of $P_{Y(n)}(z)$ that the expected value of $Y(n)$, $EY(n)$, is given by

$$EY(n) = \frac{2n}{n+1}$$

.

Proposition 1.2.3. *Let $VarY(n)$ denote the variance of $Y(n)$. Then*

$$VarY(n) = \frac{(n-1)(n^2-4n-3)}{(n+1)^2(2n+3)}$$

.

Hence, we observe that $EY(n) \to 2$ as $n \to \infty$ and $VarY(n) \to \frac{1}{2}$ as $n \to \infty$. This observation leads us to the following question.

Question 1. *What is the limiting distribution of $Y(n)$?*

1.3 ANALOGUES OF CATALAN/BINOMIAL IDENTITIES

In this section, we are concerned with comparing the function $T(z)$ with the famous Catalan function

$$C(z) = 1 + zC^2(z) = \sum_{n=0}^{\infty} \frac{1}{n+1} \binom{2n}{n} z^n$$

At this point, the most immediate similarity between these functions is the expression of their generating functions. For this reason and others, we consider $T(z)$ combinatorially analogous to $C(z)$. As two illustrations of this, we have $C(z)$ counting Dyck paths of length $2n$ and rooted planar trees with n edges;

analogously, $T(z)$ counting the objects described in §0.1. Deutsch and Shapiro proved severaly basic identities involving $C(z)$ and the generating function for the central binomial coefficients,

$$B(z) = \frac{1}{\sqrt{1-4z}} = \sum \binom{2n}{n} z^n$$

. We would like to present combinatorial proofs of two analogous identities for $T(z)$. First, we introduce $M(z) = \sum \binom{3n}{n} z^n$, the analogue of $B(z)$.

Proposition 1.3.1. *The function* $M(z) = \sum \binom{3n}{n} z^n$ *counts*

(i.) *paths in* $Z \times Z$ *starting at* $(0,0)$ *and ending at* $(3n,0)$ *using steps in* $\{(1,1),(1,-2)\}$, *and*

(ii.) *even trees with* $2n$ *edges and one node (not the root) colored red.*

We want to exploit the observation that $B(z)$ is to $C(z)$ as $M(z)$ is to $T(z)$. Note the following identities:

(1) $B(z)C^s(z) = \sum_{n=0}^{\infty} \binom{2n+s}{n} z^n$

(2) $B(z) = 1 + 2zC(z)B(z)$

Proposition 1.3.2.

(i.) $M(z) = 1 + 3zT^2(z)M(z)$

(ii.) $M(z)T^s(z) = \sum_{n=0}^{\infty} \binom{3n+s}{n} z^n$

1.4 AREA UNDER GENERALIZED DYCK PATHS

Let $\mathcal{S}_n$ denote the set of all generalized Dyck paths in $\mathcal{T}_{n,0}$ which touch the x-axis at most twice. We will refer to these paths as strict generalized Dyck paths. Let a_n^s (resp. a_n^t) denote the total area of the region bounded by the paths of $\mathcal{S}_n$ (resp. $\mathcal{T}_{n,0}$) and the x-axis. Our goal is to find a generating function or recurrence for the sequence $\{a_n^s\}_{n=0}^{\infty}$ (resp. $\{a_n^t\}_{n=0}^{\infty}$).

To approach this problem, we consider $h_{n,k}^s$ (resp. $h_{n,k}^t$), the number of lattice points in $\mathcal{S}_n \cap Z \times Z$ (resp. $\mathcal{T}_{n,0} \cap Z \times Z$), which have height k.

Theorem 1.4.1. *Let* $H_k^s(z) = \sum_{n=0}^{\infty} h_{n,k}^s z^n$ *and* $H_k^t(z) = \sum_{n=0}^{\infty} h_{n,k}^t z^n$.

(i.)

$$H_k^s(z) = \sum_{j=0}^{\lfloor \frac{k}{2} \rfloor} \binom{k-j}{j} (zT^3(z))^{k-j} = (zT^3(z))^k \frac{\frac{z^2}{1-z} T^3(\frac{z^2}{1-z})}{z^2 T^3(\frac{z^2}{1-z}) - 1}$$

(ii.)

$$H_k^t(z) = \sum_{j=0}^{\lfloor \frac{k}{2} \rfloor} \binom{k-j}{j} (zT^3(z))^{k-j} T^2(z) = (zT^3(z))^{k+2} \frac{\frac{z^2}{1-z} T^3(\frac{z^2}{1-z})}{z^2 T^3(\frac{z^2}{1-z}) - 1}$$

Currently, we do not have a "nice" closed form for the generating functions above. We would like to use these generating functions to form a managable Riordan or Riordan-like matrix. Given such a matrix, we could multiply it by the Riordan vector $(\frac{z}{1-z})$, the product being the vectors $(a_n^s)_{n=0}^{\infty}$ and $(a_n^t)_{n=0}^{\infty}$.

A recent observation regarding weighted ternary trees has led to the following conjecture.

Conjecture 1. *Let $\mathcal{W}_n$ denote the set of all ternary trees with $3n$ edges and $\mathcal{T}_n$ denote the set of all ternary paths of length n. Then $\forall n \in Z^+$,*

$$\sum_{W\in\mathcal{W}_n}\omega(W)=\sum_{T\in\mathcal{T}_n}a(T)$$

where $a(T)$ denotes the area of the region bounded by T and $\omega(W)$ denotes a special weight.

Question 2. *Can we get a closed form of the generating function for area under generalized Dyck paths?*

§2 STRUCTURE IN THE RIORDAN GROUP

2.1 ELEMENTS OF ORDER 2*

An element R of the Riordan group is said to have pseudo order two, or order $2*$, if RM has order two, where $M=(1,-z)$. A prime example of an element of order $2*$ is the generalized Pascal's triangle matrix, $\Psi_b=\left(\frac{1}{1-2bz},\frac{-z}{1-2bz}\right)$.

We would like to find combinatorial interpretations of such elements of order $2*$. We know for example that elements of the form $BMB^{-1}M$ have order $2*$. One might ask, if R has order $2*$, is it the case that $R=BMB^{-1}M$ for some B? If so, what is the combinatorial relationship between R and B? We may begin to answer this question by considering order $2*$ matrices of a special type.

Theorem 2.1.1. *Let*

$$\Theta_{b,\lambda,\epsilon,\delta}=\left(\frac{1+\frac{\epsilon z}{1-bz}C\left(\frac{\lambda z^2}{(1-bz)^2}\right)-\frac{\delta z^2}{(1-bz)^2}C^2\left(\frac{\lambda z^2}{(1-bz)^2}\right)}{1-\frac{\epsilon z}{1-bz}C\left(\frac{\lambda z^2}{(1-bz)^2}\right)-\frac{\delta z^2}{(1-bz)^2}C^2\left(\frac{\lambda z^2}{(1-bz)^2}\right)}\cdot\frac{1}{1-2bz},\frac{-z}{1-2bz}\right)$$

Then

(i.) $\Theta_{b,\lambda,\epsilon,\delta}$ has order 2.*

(ii.) There exists a Riordan matrix B such that $\Theta_{b,\lambda,\epsilon,\delta}=BMB^{-1}M$.

Corollary 2.1.2. $\Psi_b=AMA^{-1}M$ *where*

$$A=\left(\frac{2n}{2n-k-(2n-k)bz+k\sqrt{1-2bz+(b^2-4n)z^2}},\frac{1-bz-\sqrt{1-2bz+(b^2-4n)z^2}}{2nz}\right)$$

for some n and k.

Corollary 2.1.3. $\Theta_{b,\lambda,\epsilon,0} = \left(\frac{(2\lambda-\epsilon b)z+\epsilon-\epsilon\sqrt{1-2bz+(b^2-4\lambda)z^2}}{(2\lambda+\epsilon b)z-\epsilon+\epsilon\sqrt{1-2bz+(b^2-4\lambda)z^2}}\frac{1}{1-2bz}, \frac{-z}{1-2bz}\right) = BMB^{-1}M$, *where*

$$B = \left(\frac{2}{1-(2\epsilon+b)z+\sqrt{1-2bz+(b^2-4\lambda)z^2}}, \frac{1-bz-\sqrt{1-2bz+(b^2-4\lambda)z^2}}{2\lambda z}\right)$$

.

There are certainly other elements of order 2* which are not of the form in Theorem 2.1.1. One example is $\left(1+zm(z), z+z^2m(z)\right)$, where $m(z)$ is the generating function for the Motzkin numbers. This particular matrix counts directed animals. Can anything be said about these other matrices?

Question 3. *Can Theorem 2.1.1. be extended to other Riordan matrices of order 2*?*

REFERENCES

[1] L. Comtet, *Advanced Combinatorics*, translated, D. Reidel Publ. Co., 1974.

[2] E. Deutsch, S. Feretic and M. Noy, Diagonally Convex Directed Polyminoes and Even Trees: A Bijection and Related Issues, preprint.

[3] E. Deutsch and L. Shapiro, Seventeen Catalan Identities, *Bulletin of Institute of Combinatorics and its Applications*, to appear.

[4] E. Deutsch, Correspondence (1999).

[5] W. Feller, *An Introduction to Probability Theory and its Applications, Vol. 1*, Wiley and Sons, 1957.

[6] R. Graham, D. Knuth and O. Patashnik, *Concrete Mathematics*, 2nd ed., Addison-Wesley, 1994.

[7] P. Peart and L. Woodson, Triple Factorization of Some Riordan Matrices, *Fibonacci Quarterly* **31**, no. 2, (1993), 121-128.

[8] E. Pergola, R. Pinzani, S. Rinaldi and R. Sulanke, A Bijective Approach to the Area of Generalized Motzkin Paths, preprint (2000).

[9] L. Shapiro, S. Getu, W. Woan and L. Woodson, The Riordan Group, *Discrete Applied Mathematics* **34** (1991), 229-239.

[10] R. Stanley, *Enumerative Combinatorics, Vol. 2*, Cambridge University Press, 1999.